SCIENCE CHINA
Life Sciences

• COVER ARTICLE •

February 2011 Vol.54 No.2: 101–111
doi: 10.1007/s11427-011-4137-2

Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly

FENG Wei[1,2], PAN LiFeng[1] & ZHANG MingJie[1*]

[1]*Division of Life Science, Molecular Neuroscience Center, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China;*
[2]*National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China*

Received May 7, 2010; accepted July 7, 2010

NMR spectroscopy and X-ray crystallography are two premium methods for determining the atomic structures of macro-biomolecular complexes. Each method has unique strengths and weaknesses. While the two techniques are highly complementary, they have generally been used separately to address the structure and functions of biomolecular complexes. In this review, we emphasize that the combination of NMR spectroscopy and X-ray crystallography offers unique power for elucidating the structures of complicated protein assemblies. We demonstrate, using several recent examples from our own laboratory, that the exquisite sensitivity of NMR spectroscopy in detecting the conformational properties of individual atoms in proteins and their complexes, without any prior knowledge of conformation, is highly valuable for obtaining the high quality crystals necessary for structure determination by X-ray crystallography. Thus NMR spectroscopy, in addition to answering many unique structural biology questions that can be addressed specifically by that technique, can be exceedingly powerful in modern structural biology when combined with other techniques including X-ray crystallography and cryo-electron microscopy.

NMR spectroscopy, X-ray crystallography, structural biology, protein complex assembly

Citation: Feng W, Pan L F, Zhang M J. Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly. Sci China Life Sci, 2011, 54: 101−111, doi: 10.1007/s11427-011-4137-2

Since the first protein structure was determined in the middle of last century, structural biology has had a critical role in driving the development of modern biology, chiefly by providing atomic level resolution of individual as well as complexed forms of biological molecules. X-ray crystallography and NMR spectroscopy are the two main methods that have been used for deriving high resolution atomic pictures of biomolecules, though cryo-electron microscopy is gaining momentum in this research direction. X-ray crystallography was first developed to determine biomolecular structures based on X-ray diffraction of single crystals [1,2] (Figure 1B). However, many biomolecules are difficult to crystallize, or even cannot be crystallized, because of factors such as intrinsically disordered regions and the heterogeneous conformational properties of samples. Consequently, the most critical step for X-ray-based structure determination has been and will continue to be obtaining high quality crystals of samples of interest. By contrast, NMR spectroscopy can determine biomolecular structures in solution or in their solid state without the need to obtain crystals [3–8]. NMR spectroscopy has been particularly useful for defining the conformation of proteins with low molecular weights as well as proteins with large intrinsically disordered regions [4,5,9–11]. Nonetheless, inherent technical deficiencies as well as practical limitations of NMR spectroscopy limit the wider applications of the technique to

*Corresponding author (email: mzhang@ust.hk)

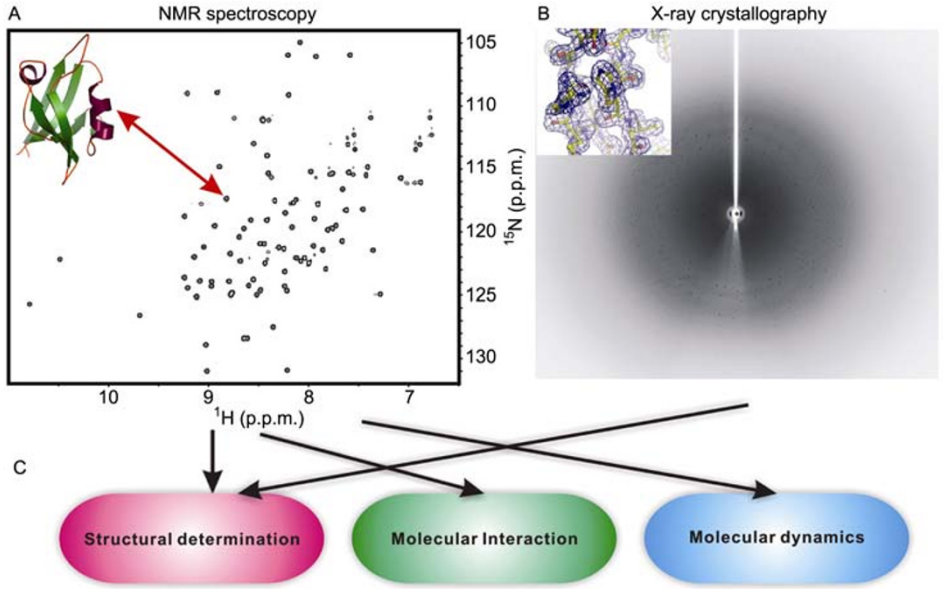

Figure 1 NMR spectroscopy and X-ray crystallography are two premium methods of structural biology. A, A typical ^{1}H-^{15}N HSQC spectrum of a protein. One cross peak in the spectrum (indicated by the arrow) represents one amino acid in the protein. B, An example X-ray diffraction map of a protein. The diffraction data can be converted into the electron density map of each amino acid in the protein. C, Understanding of the functions of biomolecules requires various information including their 3-D structures, detailed interactions, as well as dynamic properties. NMR spectroscopy can be used to determine the structures, dissect molecular interactions, and investigate the dynamics of proteins. X-ray crystallography is particularly powerful for determination of the atomic structures of highly complicated biological systems.

determine high quality structures of proteins with large molecular weights. Thus, NMR spectroscopy and X-ray crystallography are generally regarded as complementary techniques for the structure determination of biomolecules [4,12,13].

Compared with X-ray crystallography, NMR spectroscopy has unique features. The capability of NMR spectroscopy for detecting individual atoms in terms of chemical shifts, coupled with the exquisite sensitivity of chemical shifts of atoms to their conformational environment, makes it the most powerful of all spectroscopic techniques. From the structural biology viewpoint, NMR spectroscopy is the only method that allows direct monitoring of the conformational properties of a protein or its complex at the individual amino acid level (e.g. via recording ^{1}H, ^{15}N and ^{13}C chemical shifts; Figure 1A), without any prior conformational knowledge of the protein. In addition, NMR spectroscopy allows monitoring of conformational changes of a protein or its complex upon the introduction of changes to the sample condition. These two features of NMR spectroscopy render the technology exceedingly powerful when compared with X-ray crystallography, as crystallization processes are largely based on trial-and-error methods. Finally, recent developments of NMR technology (cryogenic detection system combined with ultra-high magnetic field) have essentially removed the major limitation of the technique which is its intrinsically low sensitivity. It is now possible to record a good quality NMR spectrum (e.g. ^{1}H-^{15}N HSQC spectrum) of a protein sample at a concentration as low as 10 μmol L^{-1} within a few hours. Moreover, it should be emphasized that there is essentially no molecular size limit for recording chemical shift information of proteins or their complexes by NMR spectroscopy.

NMR spectroscopy has been highly successful in the determination of biomolecular structures [3,14–16]. As the chemical shift of each atom is highly sensitive to changes induced by environmental perturbations such as target binding, NMR spectroscopy has also been used to investigate biomolecular interactions since its first applications in biology [17–19]. Furthermore, NMR spectroscopy is uniquely suited for, and has been exceedingly powerful in investigating dynamic properties of biomolecules [20–24]. The three-dimensional structure, interaction with its binding partners, and dynamic properties together provide a more complete understanding of cellular functions of a biomolecule (Figure 1C). NMR spectroscopy is probably the only technique that alone can provide atomic details in all three of these areas, and thus is recognized as an indispensable method in modern structural biology [25]. Because of space limitations, we limit this review to the combination of NMR spectroscopy and X-ray crystallography for the study of the structure and function of proteins and protein complex assemblies.

NMR spectroscopy and X-ray crystallography are highly complementary in terms of structural determinations, and in the past have been generally used separately to investigate biomolecular complexes. Recently, NMR spectroscopy has been frequently combined with X-ray crystallography to dissect the structures of complicated biomolecular complexes. The sensitivity and ultra-high resolution of NMR

chemical shifts make NMR spectroscopy a unique tool for investigating conformational properties of biomolecules at the atomic level without any prior structural information (Figure 1A). More importantly, the quality of the NMR spectrum (i.e. peak dispersion pattern, line-widths, and homogeneity) is directly related to the conformational property of the sample being investigated. Such NMR-derived information can be very valuable for obtaining high quality crystals of the molecules of interest. NMR spectroscopy can be used for sample conditioning, and tracing changes in protein behavior upon the addition of additives, and thus nicely fills the gap in the "all-or-nothing" mode of crystallization screening by X-ray crystallography. An increasing number of successful examples have demonstrated the power of NMR spectroscopy when combined with X-ray crystallography, in answering fundamental biological questions by providing high resolution structures of complicated protein assemblies [26–29].

In this review, we use several recent examples from our own laboratory to illustrate the power of the combined use of NMR spectroscopy and X-ray crystallography in solving structures of protein complexes involved in neuronal signaling and cell polarity. We focus on how we integrate information derived from NMR spectroscopy, biochemistry and crystallization screening, in an iterative manner, to obtain high quality samples for solving high resolution structures. However, we do not dwell on the biological details of each system described here, as readers can refer to the relevant papers for details. We hope that the review will provide helpful information on how to combine NMR spectroscopy and X-ray crystallography for structural studies of complicated biomolecular systems. We also wish to encourage more scientists to actively participate in using these two premium structural biology methods, in view of the major equipment build-ups under the "National Protein Science Research Infrastructure" scheme in China.

1 L27 domain-mediated supra-molecular complex assembly dissected by NMR spectroscopy

L27 domain, originally indentified in scaffold proteins Lin-2 and Lin-7, is a novel protein interaction module that mediates the assembly of supra-molecular scaffold protein complexes for membrane trafficking, polarized protein distributions, and cell-cell contacts [30]. To elucidate the structural basis of L27 domain-mediated supra-molecular assembly, we selected one pair of L27 domain complexes formed between SAP97 L27 ($L27_{SAP97}$) and mLIN2 L27N ($L27N_{mLIN2}$) to begin our study (Figure 2A). As NMR spectroscopy has always been used to monitor sample conditions and to check preliminary structural information (e.g. the proper folding indicated by specific peak dispersion patterns), we purified the isolated $L27_{SAP97}$ and $L27N_{mLIN2}$ and recorded their NMR spectra. The ^{1}H-^{15}N HSQC spectra of the isolated L27 domains demonstrated that neither of them forms a defined conformation, as indicated by severe peak broadening or the narrowly dispersed peaks that are indicative of unfolded proteins (Figures 2C1 and C2). Interestingly, when we mixed ^{15}N-labelled $L27_{SAP97}$ with unlabelled $L27N_{mLIN2}$ or *vice versa*, sharp and well-dispersed peaks appeared, indicating that the L27 domain complex (containing two related L27 domains) forms a more compact structural unit (Figures 2C3 and C4). Thus NMR spectroscopy was used to detect the specific binding of a pair of cognate L27 domains in solution, and to monitor conformational transitions of the L27 domains from the unfolded or partially unfolded states to the much better folded state upon complex formation (Figures 2C1 and C3, and Figures 2C2 and C4). To aid structural determination by the NMR-based method, we co-expressed the two L27 domains. Although the HSQC spectrum of the co-expressed sample confirmed the well-defined structural assembly of the L27 domain complex, the spectra were still not optimal for NMR structure determination (i.e. missing peaks, and very low intensities of some peaks; Figure 2C5).

To determine the complex structure we next screened a large number of sample conditions, and each step of the changes was monitored by recording an ^{1}H-^{15}N HSQC spectrum [31]. To ensure 1∶1 stoichiometry of each domain in the complex, we fused the two L27 domains into a single chain using a flexible linker. The NMR spectrum of the fused sample was of excellent quality (well-dispersed peaks with sharp line-width) and thus suitable for NMR-based structural determination (Figure 2C6). The solution structure of $L27_{SAP97}$/$L27N_{mLIN2}$ complex was determined by conventional multidimensional NMR spectroscopy. The structure revealed that the cognate L27 domain complex forms a tetramer in solution. Each L27 domain contains three helices; the first two helices (αA and αB) of each domain form a dimeric helical bundle at two opposite sites, and the last helix (αC) forms the central helical bundle to assemble two dimers into a symmetric tetramer (Figure 2B). Since each L27 domain complex is a tetramer, the cognate complexes between the "type A" orphan L27 domain and the "type B" tandem L27 domains can further assemble into a multimer that we termed supra-scaffold assemblies for very large signaling complex assembly in polarized cells [31]. The story above illustrates the iterative use of NMR spectroscopy in sample conditioning for NMR-based structural determination. We now switch to the combined use of NMR spectroscopy and X-ray crystallography in dissecting the assembly mechanism of protein complexes.

2 Myosin VI CBD/Dab2 complex structure determination: NMR leads the way

Myosin VI is the only known myosin motor that transports cargoes towards the minus end of actin-filaments [32,33].

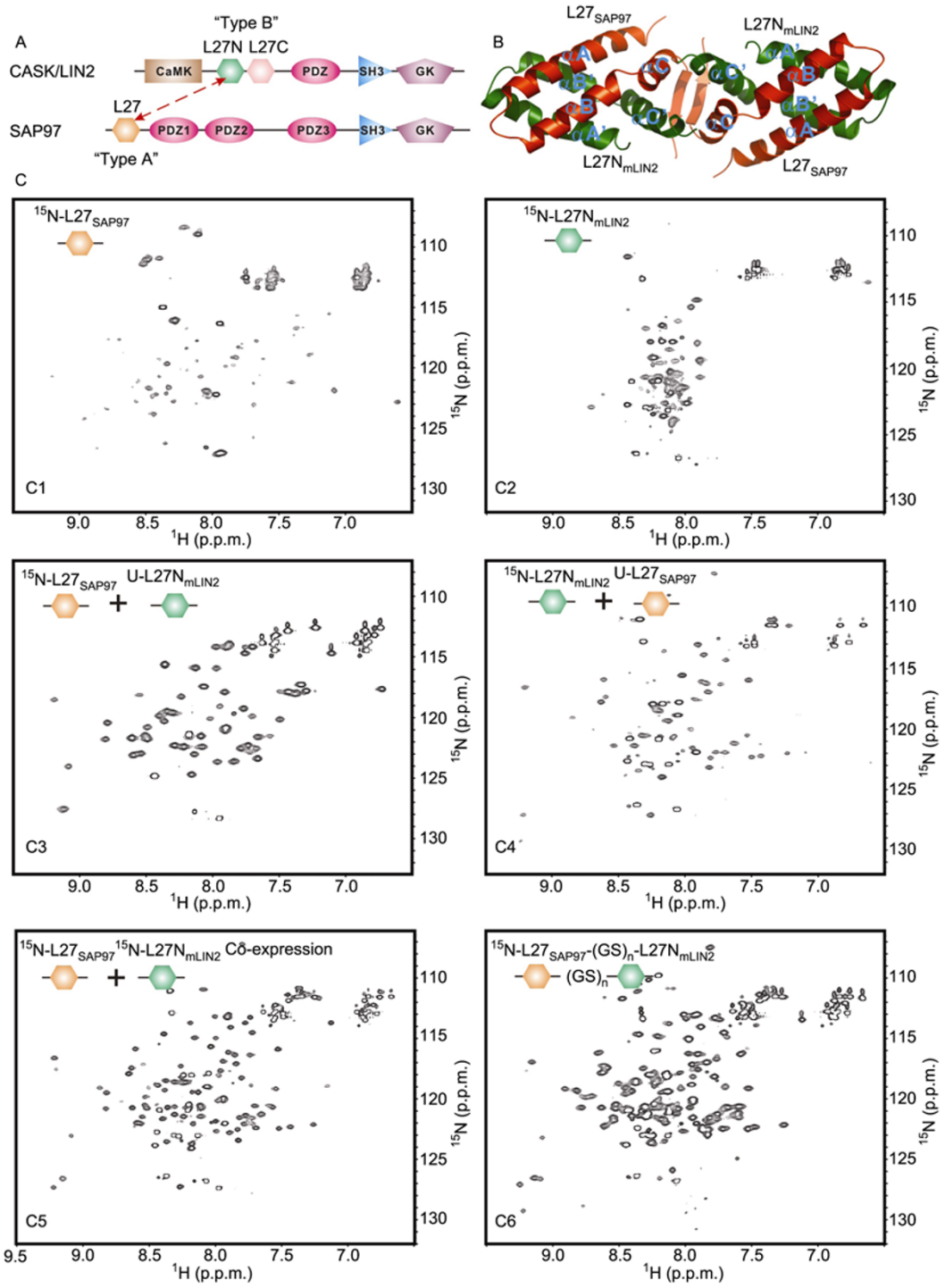

Figure 2 Dissection of the L27 domain-mediated supra-molecular complex assembly by NMR spectroscopy. A, The domain organization of CASK/LIN2 and SAP97. CASK contains the two "Type A" L27 domains arranged in tandem, namely L27N and L27C; SAP97 contains a single "Type B" L27 domain. B, A ribbon diagram representation of the tetrameric $L27_{SAP97}$/$L27N_{mLIN2}$ complex determined by NMR spectroscopy. $L27_{SAP97}$ and $L27N_{mLIN2}$ are colored orange and green, respectively. C, ^{1}H-^{15}N HSQC spectra of the isolated $L27_{SAP97}$ (C1), the isolated $L27N_{mLIN2}$ (C2), the $L27_{SAP97}$/$L27N_{mLIN2}$ complex prepared by mixing the two domains (C3-C4), the $L27_{SAP97}$/$L27N_{mLIN2}$ complex prepared by co-expression of the domains (C5), and the single-chain-fused $L27_{SAP97}$/$L27N_{mLIN2}$ complex with a $(GS)_n$ linker (C6).

Myosin VI contains an N-terminal motor domain, a short neck region and a signal α-helix (SAH) domain in the middle, and a C-terminal cargo-binding domain (CBD) (Figure 3A). Myosin VI has been reported to exist both as a monomer for anchoring and a dimer for transportation. The dimerization of myosin VI absolutely requires both the

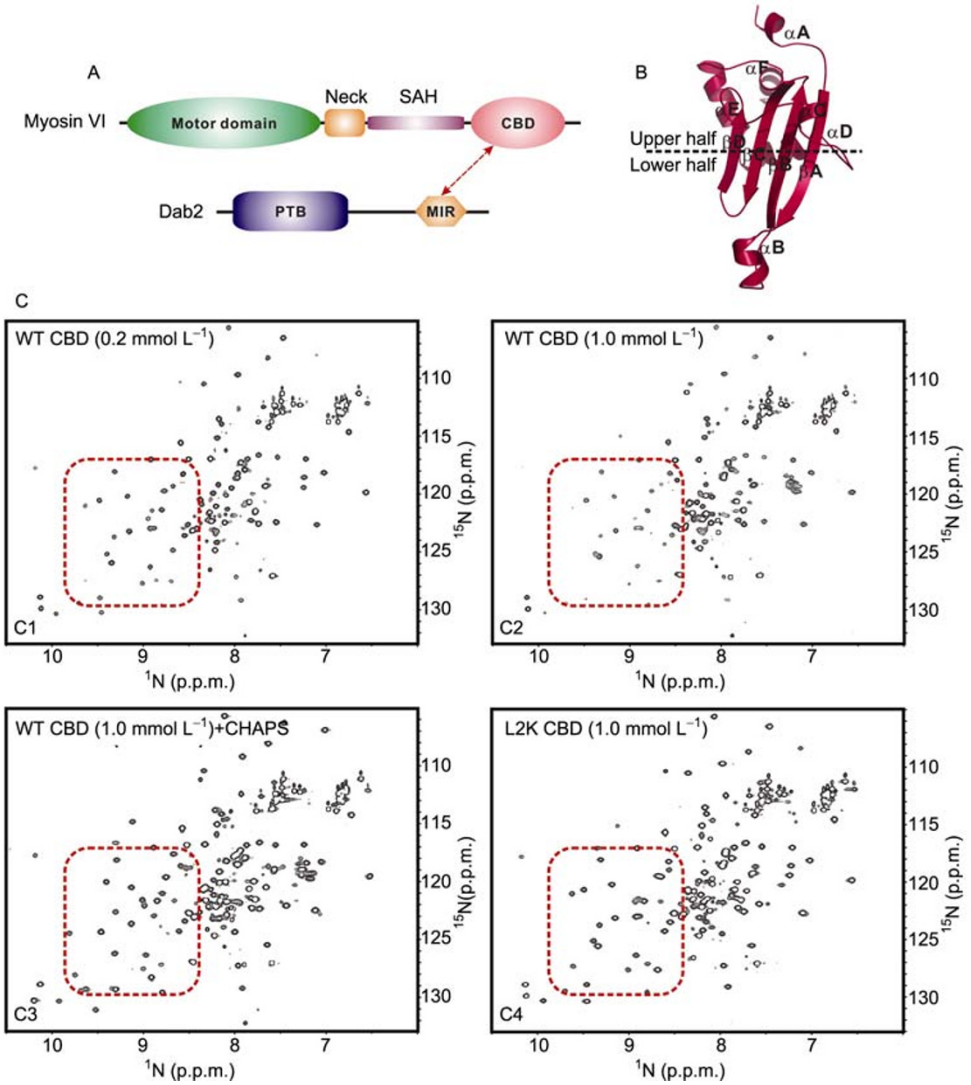

Figure 3 Structure determination of myosin VI CBD by NMR spectroscopy. A, Domain organizations of myosin VI and Dab2. Myosin VI contains an N-terminal motor domain, a short neck and a SAH domain in the middle, and a C-terminal cargo-binding domain (CDB). Dab2 contains an N-terminal PTB domain and a C-terminal myosin VI interaction region (MIR). The interaction between myosin VI and Dab2 is mediated by the CBD and the MIR. B, A ribbon diagram representation of myosin VI CBD in its apo-form. C, ^{1}H-^{15}N HSQC spectra of 0.2 mmol L^{-1} wild type (WT) CBD (C1), 1.0 mmol L^{-1} WT CBD (C2), 1.0 mmol L^{-1} WT CBD with 0.5% CHAPS (C3), and 1.0 mmol L^{-1} "L2K"-CBD (C4).

CBD and its cargo-binding [33]. To determine the structural basis of myosin VI dimerization, we first purified the CBD and used NMR spectroscopy to check the sample quality and the conformational property of the domain. The domain showed well-dispersed ^{1}H-^{15}N HSQC peaks at low sample concentrations (<0.2 mmol L^{-1}), indicating that it might be amenable for NMR-based structure determination (Figure 3C1). Unfortunately, the quality of the NMR spectra deteriorated continuously upon increase of the sample concentration (i.e. a large set of peaks severely broadened at high concentrations; Figure 3C2). The deterioration of the NMR spectra was most likely the result of non-specific aggregation of the domain at high concentrations. To overcome the aggregation problem we used the NMR-based sample condition screening, by systematically testing various additives in the sample buffer. We discovered that addition of a small amount of CHAPS, a non-ionic detergent, dramatically improved the spectral quality of the domain at a protein concentration as high as ~1.0 mmol L^{-1} (Figure 3C3). This finding suggested that hydrophobic interaction is an important factor causing the non-specific, concentration-dependent aggregation of myosin VI CBD. We next assigned the chemical shift of the domain in the presence of CHAPS, and obtained the rough topology of the domain by NMR spectroscopy. We searched for point mutations of myosin VI CBD that can disrupt the non-specific aggregation but do not alter the conformation of the domain. The site of mutation was based on the rough topology of the domain, and the overall folding of the mutants was assessed by comparing their HSQC spectra with that of the WT CBD at low con-

centration (Figure 3C1). One of the point mutations (L2K mutation with Leu1209 substituted with Lys) showed an excellent spectrum at very high concentration of the CBD, and the spectrum of the mutant was essentially the same as that of WT CBD (Figure 3C4). Using this "L2K"-mutant, we determined the solution structure of myosin VI CBD by NMR spectroscopy (Figure 3B). The structure revealed that myosin VI CBD forms a monomer and adopts a novel folding with a central β-sheet formed by four strands (βA-βD) as the core of the CBD and both sides of the β-sheet are capped with α-helices (Figure 3B). Intriguingly, all the helices are packed with the upper half of the CBD β-sheet. The lower half of the β-sheet, which is hydrophobic and functions as the binding site for the Dab2 cargo, is completely exposed to the solvent (Figure 3B). The concentration-dependent aggregation of the WT CBD also occurs via the exposed hydrophobic site, as Leu1209 is at the center of this hydrophobic surface [27].

We selected one cargo adaptor Dab2 for investigating motor/cargo interactions (Figure 3A). Dab2 is a small adaptor for clathrin-mediated endocytic vesicles. Dab2 contains an N-terminal PTB domain for binding to cell surface receptors and a C-terminal conserved region (called "MIR" for myosin VI interacting region) for binding to myosin VI (Figure 3A). The MIR was predicted to contain two helices (Figure 4A), and previous studies showed that the first helix

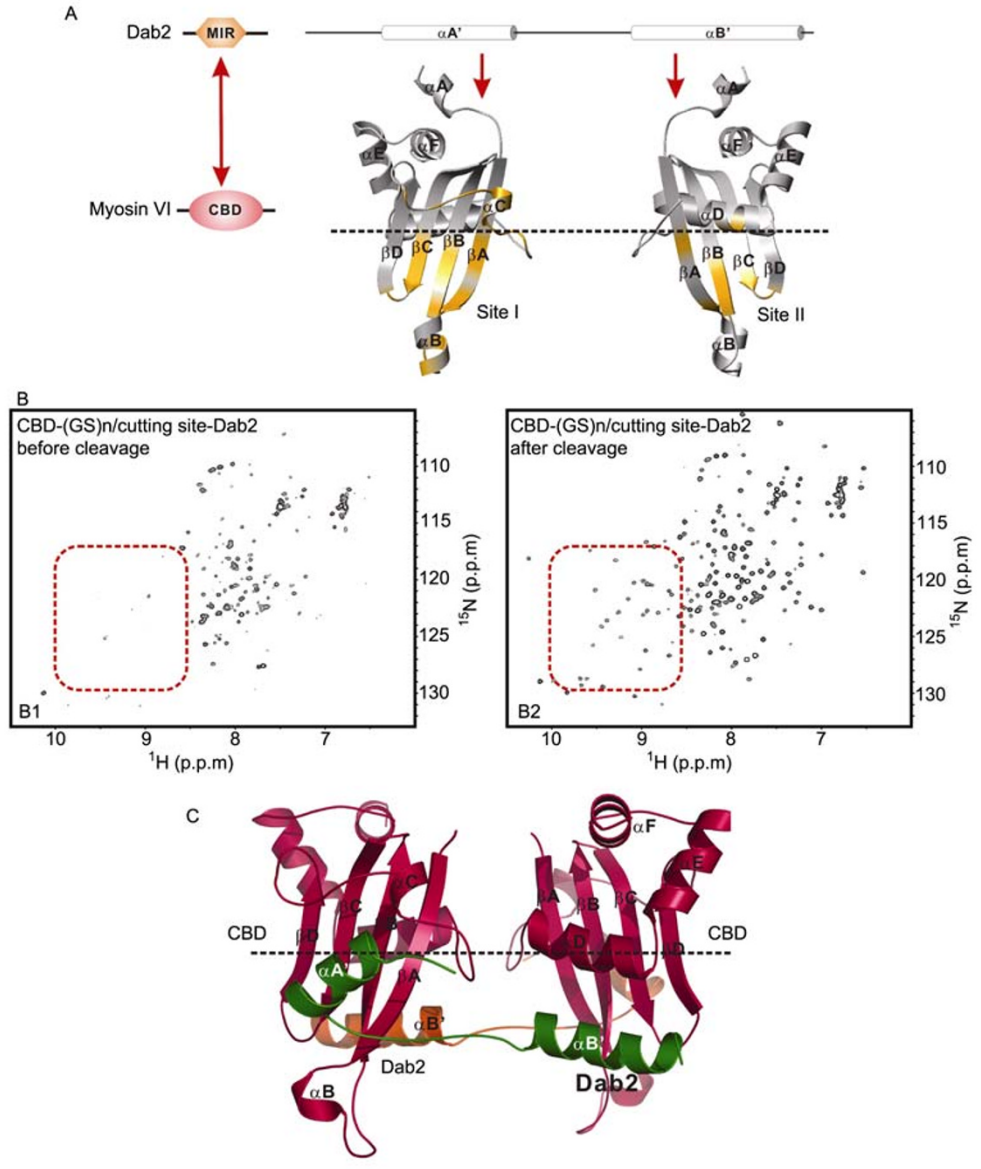

Figure 4 Structure determination of the myosin VI CBD/Dab2 MIR complex. A, The mapping of the interaction interface between myosin VI CBD and Dab2 MIR by NMR spectroscopy. The two conserved helices of Dab2 MIR bind to two exposed hydrophobic surfaces of myosin VI CBD. The first MIR helix binds to "Site I" and the second helix binds to "Site II" as indicated by red arrows. B, ^{1}H-^{15}N HSQC spectra of the single-chain-fused myosin VI CBD/Dab2 MIR complex (with a $(GS)_n$ linker and a protease cutting site), before (B1) and after (B2) cleavage. C, A ribbon diagram representation of the myosin VI CBD/Dab2 MIR complex showing MIR-induced dimerization of the CBD. The CBD and two MIRs are colored pink, green and orange, respectively.

(αA') was sufficient for its binding to myosin VI [34]. We verified the binding between myosin VI and Dab2 αA'; the two fragments bind with reasonably strong binding affinity (Kd ~0.4 μmol L^{-1}; [27]). We tried to determine the complex structure either by NMR spectroscopy or by X-ray crystallography. Under numerous conditions tested, the quality of the NMR spectra of the complex was insufficient for structure determination. Extensive trials of crystallization of the complex also failed. We decided to revisit the interaction between myosin VI CBD and Dab2 by carefully monitoring chemical shift changes in CBD upon binding to various versions of Dab2 proteins. We noticed that addition of the Dab2 fragment containing the first helix induced large chemical shift changes in myosin CBD, and these chemical shift changes were mapped to the site I pocket of the domain (Figure 4A). Unexpectedly, addition of an extended Dab2 fragment, which contains the second helix, caused further chemical shift changes in addition to the residues at site I, and the second site was termed the site II cargo binding pocket of myosin CBD (Figure 4A). The above observation revealed that the myosin VI binding region of Dab2 contains two helices instead of the previously assumed single helix [27].

Once the complete myosin VI-binding region of Dab2 was mapped, we again used NMR spectroscopy to search for the best construct/sample conditions for determining the complex structure. We tried various methods to prepare the myosin VI/Dab2 complex including co-purification, co-expression, and covalent fusion of myosin VI CBD and Dab2 MIR. Again, to ensure the stoichiometry of the CBD/Dab2 complex we fused the Dab2 MIR peptide to the C-terminus of the CBD with a linker that can be cleaved by a specific protease. The NMR spectra of the fused CBD-Dab2 protein prior to being cleaved by the protease were poor (Figure 4B1); after protease cleavage, however, the spectra were excellent (Figure 4B2). The excellent conformational homogeneity and high stability of the complex sample prompted us to screen for crystals of the complex. At the first trial with only 196 conditions using the sample prepared as described above, we obtained high quality crystals of the complex (diffracted to 2.25 Å resolution in an in-house diffractor) [27]. The crystal structure showed that the CBD/Dab2 complex forms a dimer-of-dimers assembly which is induced by the binding of the two helices in the Dab2 MIR (Figure 4C) [27]. The complex structure reveals that myosin VI (and perhaps other myosins as well) undergoes a cargo binding-mediated dimerization [27].

In the course of studying the myosin VI CBD/Dab2 complex, NMR spectroscopy was extensively used to elucidate the interactions leading to the aggregation of the isolated CBD. NMR was also used to find the optimal sample condition for both NMR- and X-ray-based structure determinations. Importantly, based on its great sensitivity in detecting binding-induced structural changes, NMR was used to discover the previously missed critical component in Dab2 that is necessary for the formation of the functional myosin VI/Dab2 complex. We emphasize that the utilization of NMR spectroscopy in sample condition screening can greatly enhance crystallization success rates. We almost never needed to set up a large number of crystal screening conditions to obtain high quality crystals in systems when NMR spectroscopy provided favorable conformational knowledge.

3 Harmonin NPDZ1/Sans SAM-PBM complex structure determination: it could not have been done without the assistance of NMR spectroscopy

Harmonin (USH1C) is a central scaffold protein involved in the organization of the USH1 protein network, which is essential to establish and maintain hair cell stereocilia [35–37]. Harmonin contains an N-terminal domain followed by two PDZ domains, a central coiled-coil region and a PDZ domain at the C-terminus (Figure 5A). Sans (USH1G) is another scaffold protein among the USH1 proteins [38]. Sans contains four N-terminal ankyrin repeats, a central region, and a C-terminal SAM domain. The extreme C-terminal tail also possesses a classic PDZ binding motif (PBM) (Figure 5A). It was reported that the first PDZ domain of Harmonin binds to the SAM domain from Sans [39], but the structural basis of such noncanonical PDZ/target interactions was unclear.

To clarify the molecular basis of the Harmonin PDZ/Sans SAM interaction, we sought to determine the complex structure. The N-terminal domain together with the first PDZ domain of Harmonin forms an NPDZ1 supramodule, which was used for further characterization of its interactions with Sans [28]. The high affinity and stoichiometric interaction between Harmonin and Sans was mapped to the N-terminal domain and PDZ1 supramodule (NPDZ1) of Harmonin and the SAM domain containing PBM of Sans (Figure 5A). In view of the high molecular weight of the complex, we decided to solve the NPDZ1/SAM-PBM complex structure by X-ray crystallography. The complex prepared from the wild type Harmonin NPDZ1 and Sans SAM-PBM formed large crystals but diffracted poorly (to only 3.6 Å resolution). Extensive trials for new crystallization conditions failed to improve the quality of the complex crystals. We again resorted to NMR spectroscopy for answers. To simplify the experimental system, we screened the conditions of NPDZ1 and SAM-PBM separately. In the NPDZ1 sample condition screening, to increase the protein expression and solubility we fused the Sans PBM to the C-terminus of the NPDZ1. The poor NMR spectrum of the wild type NPDZ1-Sans-PBM indicated that the protein is prone to non-specific aggregation (Figure 5C). We discovered by screening for various additives in the sample buffer that inclusion of high salt could markedly improve the spec-

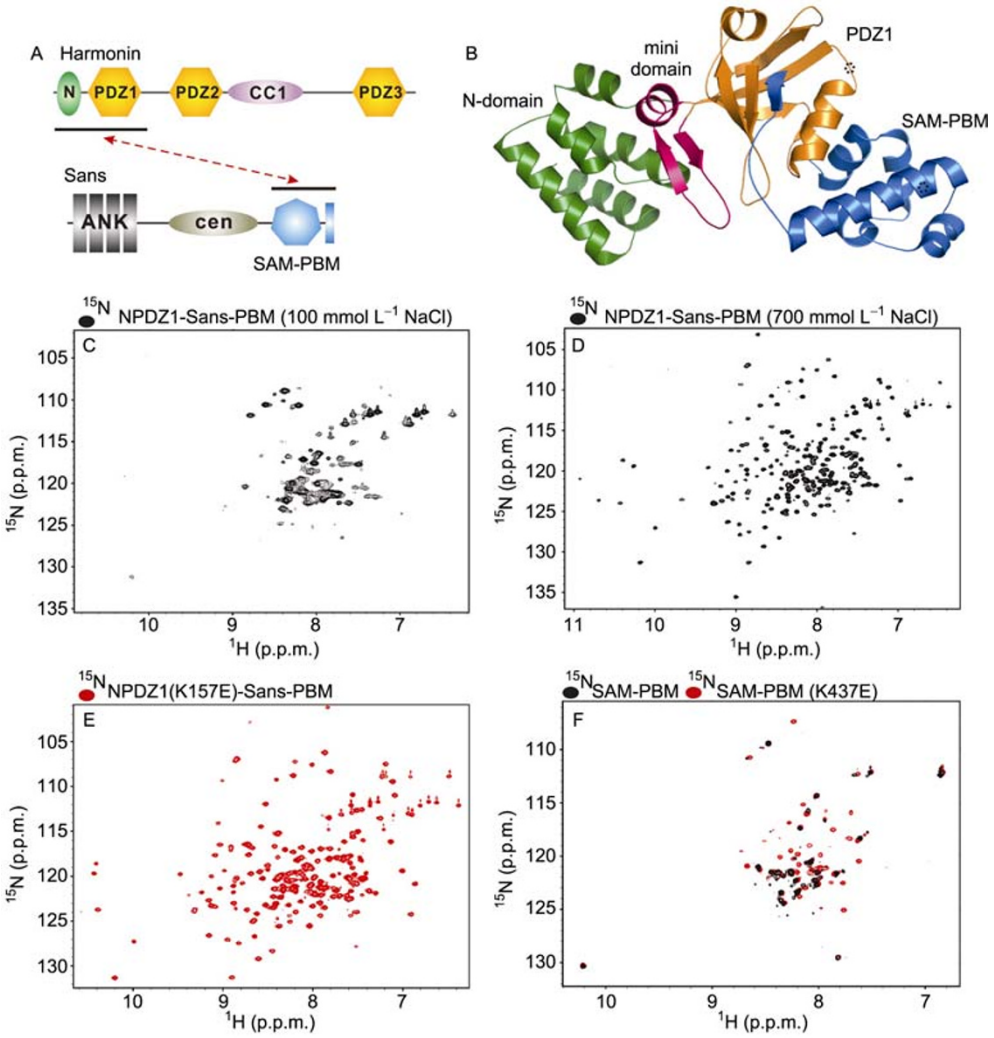

Figure 5 Structure determination of the Harmonin NPDZ1/Sans SAM-PBM complex. A, Domain organizations of Harmonin and Sans. Harmonin contains an N-terminal domain, three PDZ domains and a coiled-coil (CC1) domain in the middle. Sans contains an N-terminal ankyrin repeats (ANK) domain, a central domain (cen), and a C-terminal SAM domain with a PDZ-binding motif (PBM). B, A ribbon diagram representation of the Harmonin NPDZ1/Sans SAM-PBM complex. The N-domain, PDZ1, SAM-PBM and the mini domain are colored green, orange, blue and pink, respectively. The two point mutations identified to facilitate complex crystallization are indicated with dotted circles. C–E, ^{1}H-^{15}N HSQC spectra of NPDZ1-Sans-PBM with 100 mmol L^{-1} NaCl (C), NPDZ1-Sans-PBM with 700 mmol L^{-1} NaCl (D), and NPDZ1(K157E)-Sans-PBM with 100 mmol L^{-1} NaCl (E). F, An overlay plot of the ^{1}H-^{15}N HSQC spectra of SAM-PBM and SAM-PBM (K437E).

tral quality of the fused protein (Figure 5D), indicating that the concentration-dependent aggregation is mediated mainly by charge-charge interactions. Aided by extensive amino acid sequence analysis and modeling, we found that substitution of Lys157 by Glu eliminated the nonspecific aggregation of NPDZ1, as the NMR spectrum of the mutant in low salt buffer was highly homogenous (Figure 5E). The complex prepared using the K157E-NPDZ1 mutant and the wild-type SAM-PBM readily formed crystals that were diffracted to 3.2 Å resolution (vs 3.6 Å for the WT NPDZ1). Encouraged by this positive result, we tried to improve the sample property of SAM-PBM, as the wild type protein tends to aggregate and form various homo-oligomers in solution (Figure 5F). Again using NMR spectroscopy-based sample condition screening, we found that mutation of Lys437 on the SAM domain of Sans with Glu significantly improved the conformational homogeneity of SAM-PBM (Figure 5F). To our delight, the crystals obtained using the K157E-NPDZ1/K437E-SAM-PBM complex were diffracted to 2.3 Å. The structure of the complex was successfully determined using molecular replacement methods (Figure 5B). The structure shows that the Harmonin NPDZ1 indeed forms a supramodule with its N-domain and PDZ1 integrated by a mini-domain (a 25-residue extension C-terminal to PDZ1). Unexpectedly, the Sans SAM domain packs extensively with the αB and βE of PDZ1, forming a novel interaction mode for PDZ domains (Figure 5B). Consequently, the NPDZ1/SAM-PBM complex structure reveals a completely new PDZ domain-mediated protein/protein interaction mode [28].

In the NPDZ1/SAM-PBM complex study, NMR spectroscopy was extensively incorporated into X-ray crystallography at various stages of the project (e.g. for monitoring sample quality, identifying residues responsible for non-specific aggregation, and discovering specific mutations for preparing high diffraction quality crystals) [28]. It should be emphasized here that we have taken great care in designing point mutations of various proteins for structural studies by minimizing possible structural changes introduced by mutations. Again, NMR spectroscopy has played a crucial role in such exercises, as the technique allows us to monitor chemical shift changes introduced by mutations. For example, the two mutations used in determining the Harmonin/San complex (K157E of Harmonin NPDZ1 and K473E of Sans SAM-PBM) are both on the solvent exposed surface of each protein and away from the interface of the complex (Figure 5B).

4 Uncovering of the closed conformation of lipidated Ykt6: NMR spectroscopy provides a direction for crystallography

Ykt6 is an evolutionarily conserved SNARE protein, essential for vesicle membrane fusion at the Golgi, vacuoles and endosomes [40–43]. Ykt6 is composed of an N-terminal longin domain followed by a SNARE core, and a C-terminal double lipidation "CCAIM"-motif (Figure 6A). Biochemistry and cell biology studies of Ykt6 have shown that it exists in both membrane-bound and soluble cytosolic forms, which are controlled by lipidations (i.e. the farnesylation and the palmitoylation) [42–45]. However, the molecular basis of lipidation-mediated conformational regulation of Ykt6 was unclear, as the unlipidated full length Ykt6 is conformationally heterogeneous [44], and farnesylated Ykt6 could not be crystallized [46]. To determine the structural basis of lipidation-mediated Ykt6 membrane/cytosol cycling, we set out to study the interactions between lipids and Ykt6 using NMR spectroscopy.

Because it was suggested that the longin domain of yeast Ykt6 binds to palmitoyl-CoA (Pal-CoA), we tested for potential direct interactions between the Ykt6 longin domain and Pal-CoA using NMR spectroscopy. However, we could not detect any interactions between the Ykt6 longin domain and Pal-CoA (<40 μmol L^{-1} to avoid micelle formation) in our NMR-based assay. We screened several additional lipids for potential binding to the Ykt6 longin domain, and found that DPC (dodecylphosphocholine, with critical micelle concentration ~1.4 mmol L^{-1}) binds weakly to Ykt6 longin domain (Figure 6B) [29]. To further investigate the interaction with DPC, we next tested the binding of DPC to the full length of Ykt6 (including both the longin domain and the SNARE core).

To our surprise, the addition of stoichiometric amounts of DPC induced large chemical shift changes for many residues in Ykt6 (Figure 6C, with Kd~60 μmol L^{-1} from the DPC titration), revealing that the full-length Ykt6 binds to DPC with much higher affinity than does the longin domain alone. The HSQC spectrum of the full-length Ykt6 in the absence of DPC (Figure 6D, blue peaks) showed that for a number of residues each backbone amide displayed several peaks, indicating that the protein exists in multiple conformations exchanging at slow-to-intermediate timescales. The high sensitivity and resolution of NMR chemical shifts allowed us to observe that one of these conformations coincides with the fully open structure, as the amide peaks of this state overlap well with the corresponding amide peaks of the longin domain of the protein (Figure 6D, yellow peaks). Another conformation matches with the structure of the protein saturated with DPC (i.e. the fully closed conformation; Figure 6D, pink peaks). The rest of the conformations of Ykt6 are between the fully open and completely closed state of the protein. This NMR-based study allowed us to conclude that the unlipidated full-length Ykt6 is highly dynamic, with conformations ranging from the fully open to the completely closed states interconverting in solution.

The excellent NMR spectrum of Ykt6 in the presence of stoichiometric ratios of DPC indicates that under these conditions, the protein adopts a stable and monodisperse conformation. Guided by this information, we succeeded in obtaining high quality Ykt6 crystals by simply mixing the protein sample in the presence of ~1 mmol L^{-1} DPC with ammonium sulfate. The structure of the full-length Ykt6 in the complex with DPC was solved at 2.4 Å resolution (Figure 6E) [29]. The structure of Ykt6/DPC complex reveals that the full length Ykt6 indeed adopts a fully closed conformation with a compact globular shape, i.e. the entire SNARE core wraps around the well-defined longin domain (Figure 6E). We further showed, with the aid of cell biology data and molecular dynamics simulations, that the Ykt6/DPC complex structure faithfully represents the structure of the farnesylated Ykt6 [29]. The Ykt6/DPC complex also shows that the farnesyl group in Ykt6 functions both as a classical membrane localization signal as well as a direct activity regulator.

In the Ykt6/DPC complex study, NMR spectroscopy was instrumental in our discovery that DPC binds to Ykt6 in a stoichiometric manner. In addition, NMR was again shown to be exceedingly powerful in detecting conformational dynamics and assigning each sub-state among the conformational ensembles [29]. The examples demonstrate the power of NMR spectroscopy in elucidating structure-function relationships of complicated biological systems.

5 Concluding remarks

As a premium method for structural biology with its own unique features, NMR spectroscopy is highly advantageous to analyze complex biomolecular assemblies, characterize

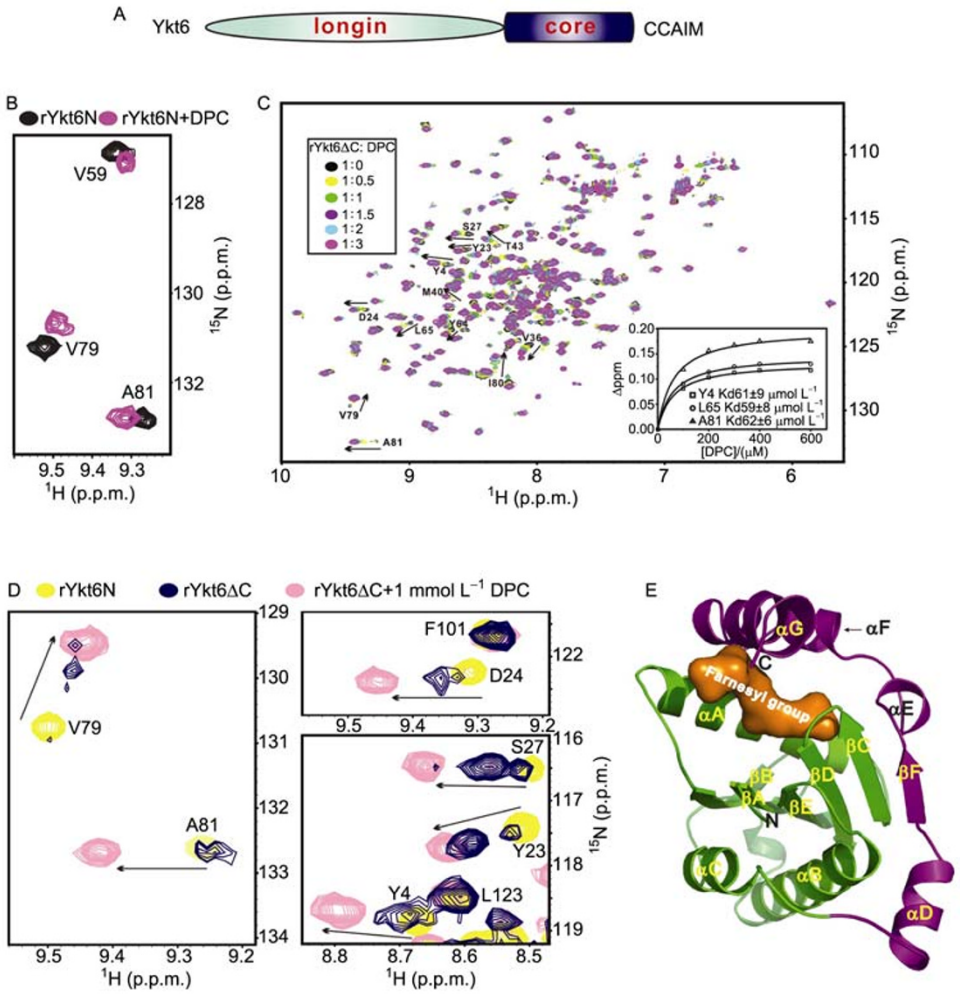

Figure 6 Lipidation-induced conformational switch of the Ykt6 SNARE. (A) Domain organization of Ykt6. Ykt6 contains an N-terminal longin domain and a SNARE core, and a C-terminal "CCAIM" lipidation motif. (B) An overlay plot of the selected region of the ^{1}H-^{15}N HSQC spectra of Ykt6N (N-terminal longin domain) and Ykt6N saturated with DPC. (C) An overlay plot of the ^{1}H-^{15}N HSQC spectra of a series of the titration points with different molar ratios of Ykt6ΔC (containing the login domain and the SNARE core without the "CCAIM" motif) to DPC. (D) An overlay plot of selected regions of the ^{1}H-^{15}N HSQC spectra of Ykt6N (colored yellow), Ykt6ΔC (colored blue) and Ykt6ΔC saturated with DPC (colored pink). (E) A ribbon diagram representation of the Ykt6/DPC complex. The entire SNARE core (colored purple) is divided into several fragments (αD-αG and βF) and tightly wraps around the longin domain (colored green). The farnesyl group (in surface representation and colored orange) is modeled into the DPC binding pocket and locks Ykt6 in a closed conformation.

protein/protein or protein/ligand interactions, monitor conformational status and sample quality of proteins, and identify suitable sample conditions for high resolution structure determination by both NMR spectroscopy and X-ray crystallography. The examples discussed in this review touched on only a small proportion of the research that can benefit by combining NMR spectroscopy with other biophysical techniques (X-ray crystallography in particular). Nonetheless, these examples strongly support our contention that the combination of NMR spectroscopy with other mainstream methods such as X-ray crystallography and cryo-electron microscopy can accelerate the process of discovery and broaden research horizons in modern structural biology.

We thank Dr. Wen Wenyu for help in preparing Figure 6. Research in our laboratory has been supported by grants from the Research Grants Council of Hong Kong to M.Z. and W.F. W.F. is now supported by the National Major Basic Research Program of China (Grant No. 2011CB910500), the National Natural Science Foundation of China (Grant No. 31070657) and the Knowledge Innovation Program of the Chinese Academy of Sciences (Grant No. KSCX2-YW-R-154). The NMR spectrometers used in our studies were funded by donations from the Hong Kong Jockey Club Charity Foundation and the Special Equipment Grant from RGC of Hong Kong (Grant No. SEG_HKUST06).

1 Kendrew J C, Bodo G, Dintzis H M, *et al*. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature, 1958, 181: 662–666
2 Muirhead H, Perutz M F. Structure of Haemoglobin. A Three-Dimen-

sional Fourier Synthesis of Reduced Human Haemoglobin at 5-5 a Resolution. Nature, 1963, 199: 633–638
3 Wuthrich K. NMR of proteins and nucleic acids. New York: John Wiley & Sons, Inc. 1986
4 Wuthrich K. Protein structure determination in solution by NMR spectroscopy. J Biol Chem, 1990, 265: 22059–22062
5 Wagner G. An account of NMR in structural biology. Nat Struct Biol, 1997, 4(Suppl): 841–844
6 Fu R, Cross T A. Solid-state nuclear magnetic resonance investigation of protein and polypeptide structure. Annu Rev Biophys Biomol Struct, 1999, 28: 235–268
7 Tycko R. Biomolecular solid state NMR: advances in structural methodology and applications to peptide and protein fibrils. Annu Rev Phys Chem, 2001, 52: 575–606
8 McDermott A E. Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. Curr Opin Struct Biol, 2004, 14: 554–561
9 Cavanagh J, Fairbrother W J, III AGP, Skeleton NJ Protein NMR spectroscopy: principles and practice. San Diego: Academic Press, 1996
10 Mittag T, Forman-Kay J D. Atomic-level characterization of disordered protein ensembles. Curr Opin Struct Biol, 2007, 17: 3–14
11 Eliezer D. Biophysical characterization of intrinsically disordered proteins. Curr Opin Struct Biol, 2009, 19: 23–30
12 MacArthur M W, Driscoll P C, Thornton J M. NMR and crystallography--complementary approaches to structure determination. Trends Biotechnol, 1994, 12: 149–153
13 Brunger A T. X-ray crystallography and NMR reveal complementary views of structure and dynamics. Nat Struct Biol, 1997, 4(Suppl): 862–865
14 Bax A. Two-dimensional NMR and protein structure. Annu Rev Biochem, 1989, 58: 223–256
15 Bax A, Grzesiek S. Methodological advances in protein NMR. Acc Chem Res, 1993, 26: 131–138
16 Kay L E, Gardner K H. Solution NMR spectroscopy beyond 25 kDa. Curr Opin Struct Biol, 1997, 7: 722–731
17 Zuiderweg E R. Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry, 2002, 41: 1–7
18 Carlomagno T. Ligand-target interactions: what can we learn from NMR? Annu Rev Biophys Biomol Struct, 2005, 34: 245–266
19 Takeuchi K, Wagner G. NMR studies of protein interactions. Curr Opin Struct Biol, 2006, 16: 109–117
20 Kay L E. Protein dynamics from NMR. Nat Struct Biol, 1998, 5(Suppl): 513–517
21 Ishima R, Torchia D A. Protein dynamics from NMR. Nat Struct Biol, 2000, 7: 740–743
22 Palmer A G 3rd. Nmr probes of molecular dynamics: overview and comparison with other techniques. Annu Rev Biophys Biomol Struct, 2001, 30: 129–155
23 Case D A. Molecular dynamics and NMR spin relaxation in proteins. Acc Chem Res, 2002, 35: 325–331
24 Korzhnev D M, Kay L E. Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. Acc Chem Res, 2008, 41: 442–451
25 Markwick P R, Malliavin T, Nilges M. Structural biology by NMR: structure, dynamics, and interactions. PLoS Comput Biol, 2008, 4: e1000168
26 Lu J, Machius M, Dulubova I, *et al*. Structural basis for a Munc13-1 homodimer to Munc13-1/RIM heterodimer switch. PLoS Biol, 2006, 4: e192
27 Yu C, Feng W, Wei Z, *et al*. Myosin VI undergoes cargo-mediated dimerization. Cell, 2009, 138: 537–548
28 Yan J, Pan L, Chen X, *et al.* The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc Natl Acad Sci USA, 2010, 107: 4040–4045
29 Wen W, Yu J, Pan L, *et al*. Lipid-Induced Conformational Switch Controls Fusion Activity of Longin Domain SNARE Ykt6. Mol Cell, 2010, 37: 383–395
30 Doerks T, Bork P, Kamberov E, *et al*. L27, a novel heterodimerization domain in receptor targeting proteins Lin-2 and Lin-7. Trends Biochem Sci, 2000, 25: 317–318
31 Feng W, Long J F, Fan J S, *et al*. The tetrameric L27 domain complex as an organization platform for supramolecular assemblies. Nat Struct Mol Biol, 2004, 11: 475–480
32 Buss F, Spudich G, Kendrick-Jones J. Myosin VI: cellular functions and motor properties. Annu Rev Cell Dev Biol, 2004, 20: 649–676
33 Sweeney H L, Houdusse A. What can myosin VI do in cells? Curr Opin Cell Biol, 2007, 19: 57–66
34 Morris S M, Arden S D, Roberts R C, *et al*. Myosin VI binds to and localises with Dab2, potentially linking receptor-mediated endocytosis and the actin cytoskeleton. Traffic, 2002, 3: 331–341
35 Reiners J, Wolfrum U. Molecular analysis of the supramolecular usher protein complex in the retina. Harmonin as the key protein of the Usher syndrome. Adv Exp Med Biol, 2006, 572: 349–353
36 Grillet N, Xiong W, Reynolds A, *et al*. Harmonin mutations cause mechanotransduction defects in cochlear hair cells. Neuron, 2009, 62: 375–387
37 Verpy E, Leibovici M, Zwaenepoel I, *et al*. A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. Nat Genet, 2000, 26: 51–55
38 Weil D, El-Amraoui A, Masmoudi S, *et al*. Usher syndrome type I G (USH1G) is caused by mutations in the gene encoding SANS, a protein that associates with the USH1C protein, harmonin. Hum Mol Genet, 2003, 12: 463–471
39 Adato A, Michel V, Kikkawa Y, *et al*. Interactions in the network of Usher syndrome type 1 proteins. Hum Mol Genet, 2005, 14: 347–356
40 Rossi V, Banfield D K, Vacca M, *et al*. Longins and their longin domains: regulated SNAREs and multifunctional SNARE regulators. Trends Biochem Sci, 2004, 29: 682–688
41 Kweon Y, Rothe A, Conibear E, *et al*. Ykt6p is a multifunctional yeast R-SNARE that is required for multiple membrane transport pathways to the vacuole. Mol Biol Cell, 2003, 14: 1868–1881
42 McNew J A, Sogaard M, Lampen N M, *et al*. Ykt6p, a prenylated SNARE essential for endoplasmic reticulum-Golgi transport. J Biol Chem, 1997, 272: 17776–17783
43 Meiringer C T, Auffarth K, Hou H, *et al*. Depalmitoylation of Ykt6 prevents its entry into the multivesicular body pathway. Traffic, 2008, 9: 1510–1521
44 Tochio H, Tsui M M, Banfield D K, *et al*. An autoinhibitory mechanism for nonsyntaxin SNARE proteins revealed by the structure of Ykt6p. Science, 2001, 293: 698–702
45 Fukasawa M, Varlamov O, Eng W S, *et al*. Localization and activity of the SNARE Ykt6 determined by its regulatory domain and palmitoylation. Proc Natl Acad Sci USA, 2004, 101: 4815–4820
46 Pylypenko O, Schonichen A, Ludwig D, *et al*. Farnesylation of the SNARE protein Ykt6 increases its stability and helical folding. J Mol Biol, 2008, 377: 1334–1345